Anticipate & Act: Integrating LLMs and Classical Planning for Efficient Task Execution in Household Environments

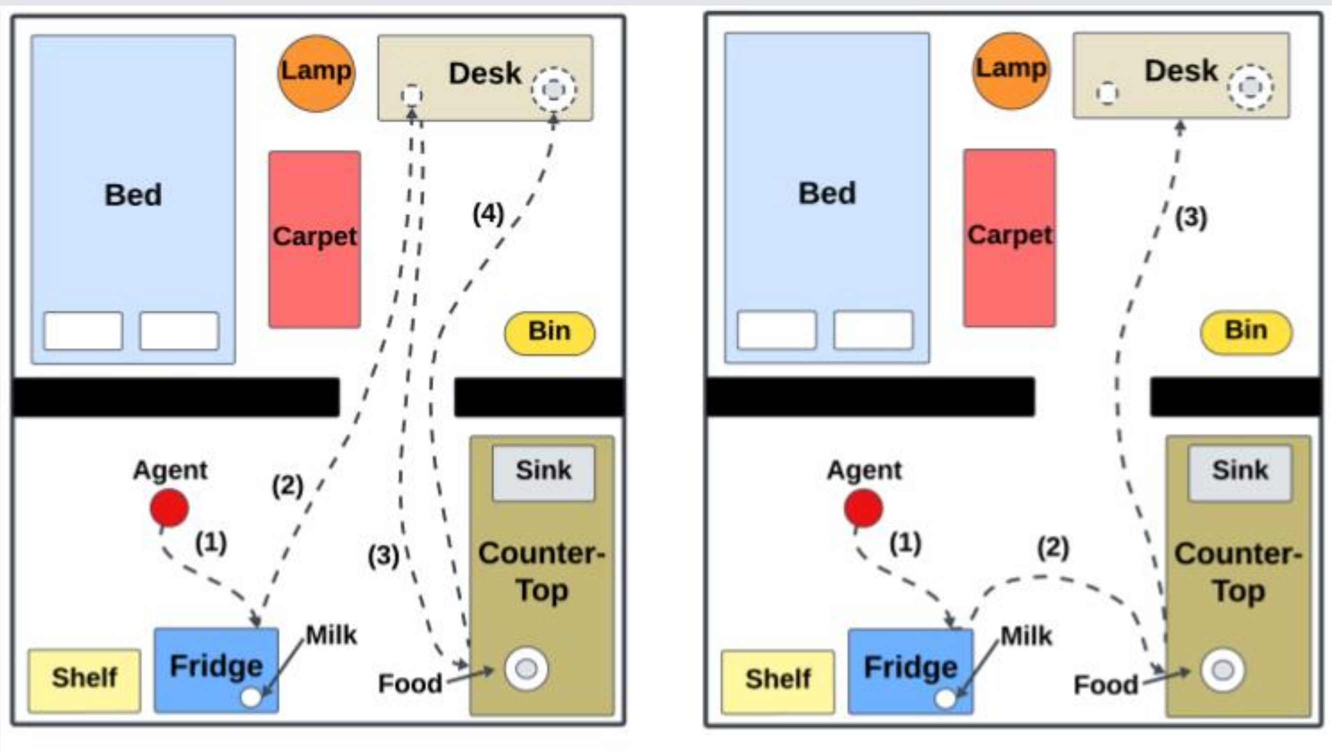
Raghav Arora[1]*, Shivam Singh[1]*, Karthik Swaminathan[1], Ahana Datta[1], Snehasis Banerjee[1,2], Brojeshwar Bhowmick[2], Krishna Murthy Jatavallabhula[3], Mohan Sridharan[4], Madhava Krishna[1]

Website: https://raraghavarora.github.io/ahsoka/
Contact: RAraghavaurora@gmail.com

[1]Robotics Research Center, IIIT Hyderabad, India
[2]TCS Research, Tata Consultancy Services, India
[3]CSAIL, Massachusetts Institute of Technology, USA
[4]IPAB, University of Edinburgh, UK

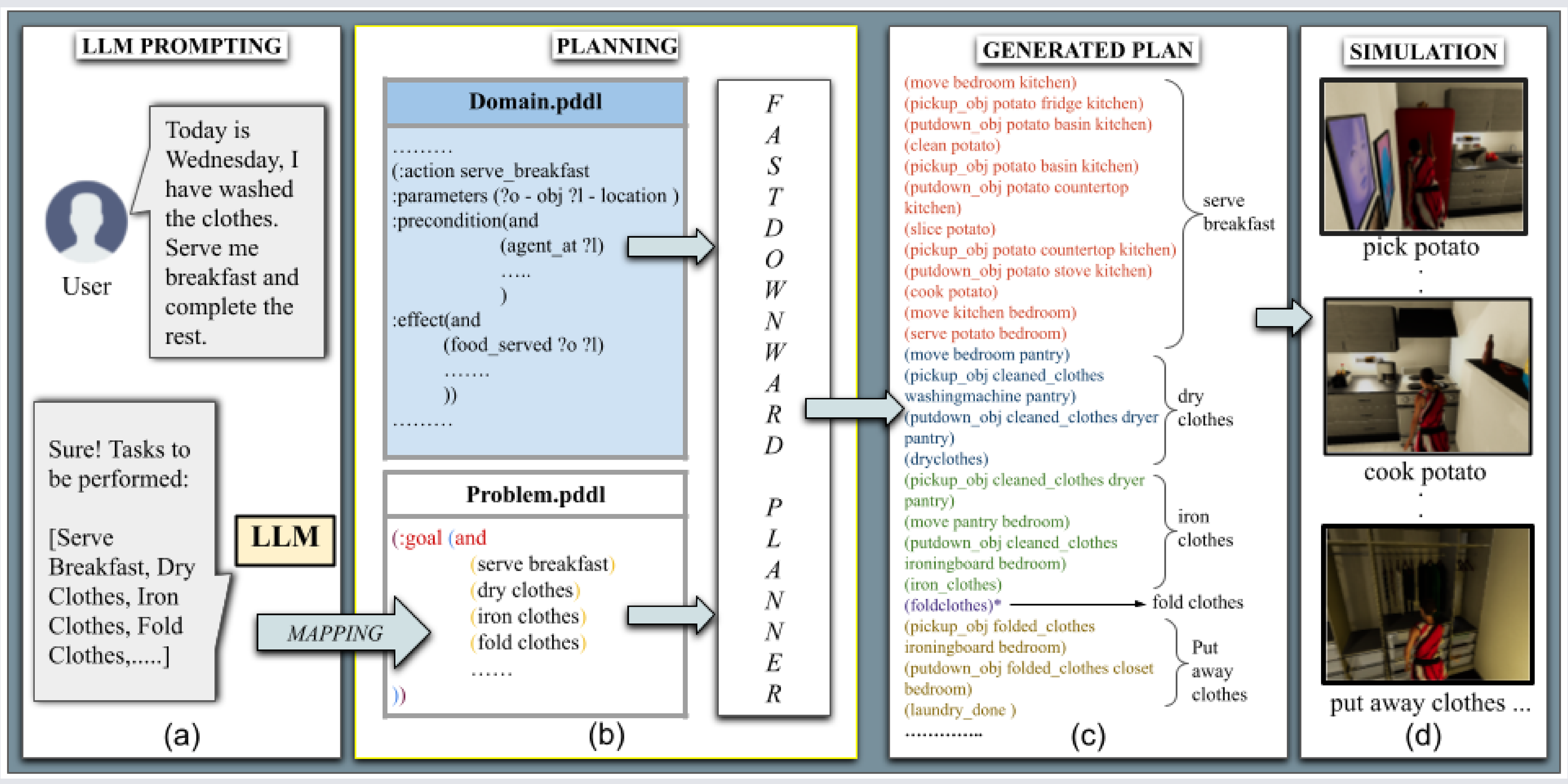
Introduction

- Assistive agents performing household tasks often compute and execute actions that accomplish one task at a time.
- Efficiency can be improved by anticipating upcoming tasks and computing an action sequence that jointly achieves these tasks.
- We use:
  - world knowledge of LLMs for high-level task anticipation
  - classical planning system to compute a sequence of finer granularity actions
  - realistic scenarios in the *VirtualHome* environment for task execution and grounding.



(a) (b)

Framework



Routine: $\mathcal{R} = \{\tau_1, \tau_2, \ldots, \tau_n\}$

$\forall\, \tau_j \in \mathcal{T}$ (known tasks)

LLM objective: predicting tasks τ_i for a routine $\mathcal{R}$

Each task τ_j requires a sequence $\{a_1^j, a_2^j, \ldots, a_k^j\}$ to be executed

Every action a_k^j has a cost c_k^j

Plan : $\pi = (a_1, \ldots, a_K)$

Planner objective : $\pi^* = (argmin)_{\pi^j} \mathcal{C}(\pi^j)$,

where $\mathcal{C}(\pi^j) = \sum_{k=0}^{K} c_k^j$

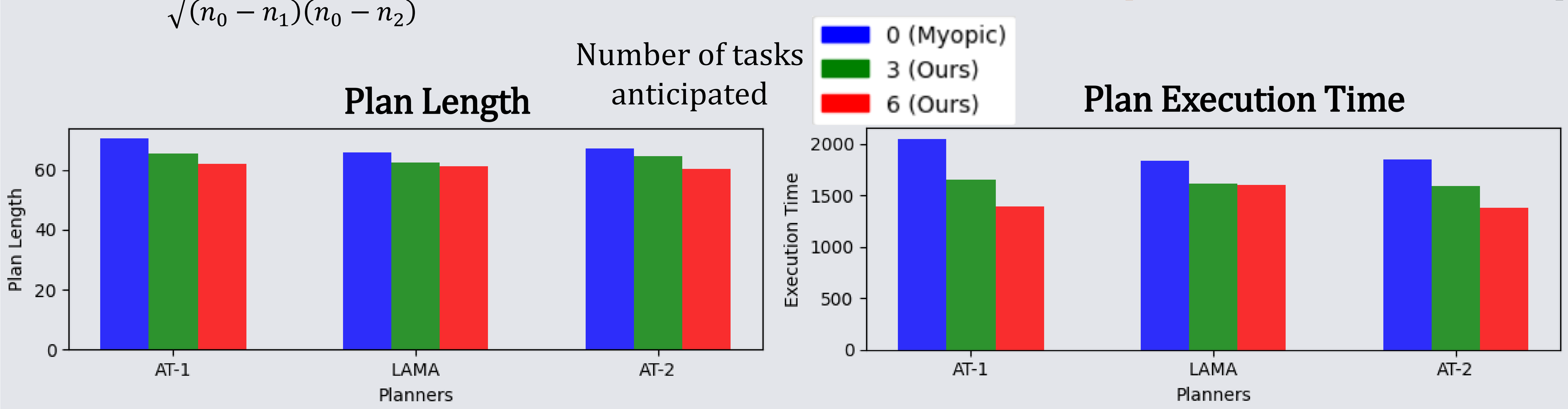
Cost c_k^j represents the time taken by the agent for execution.

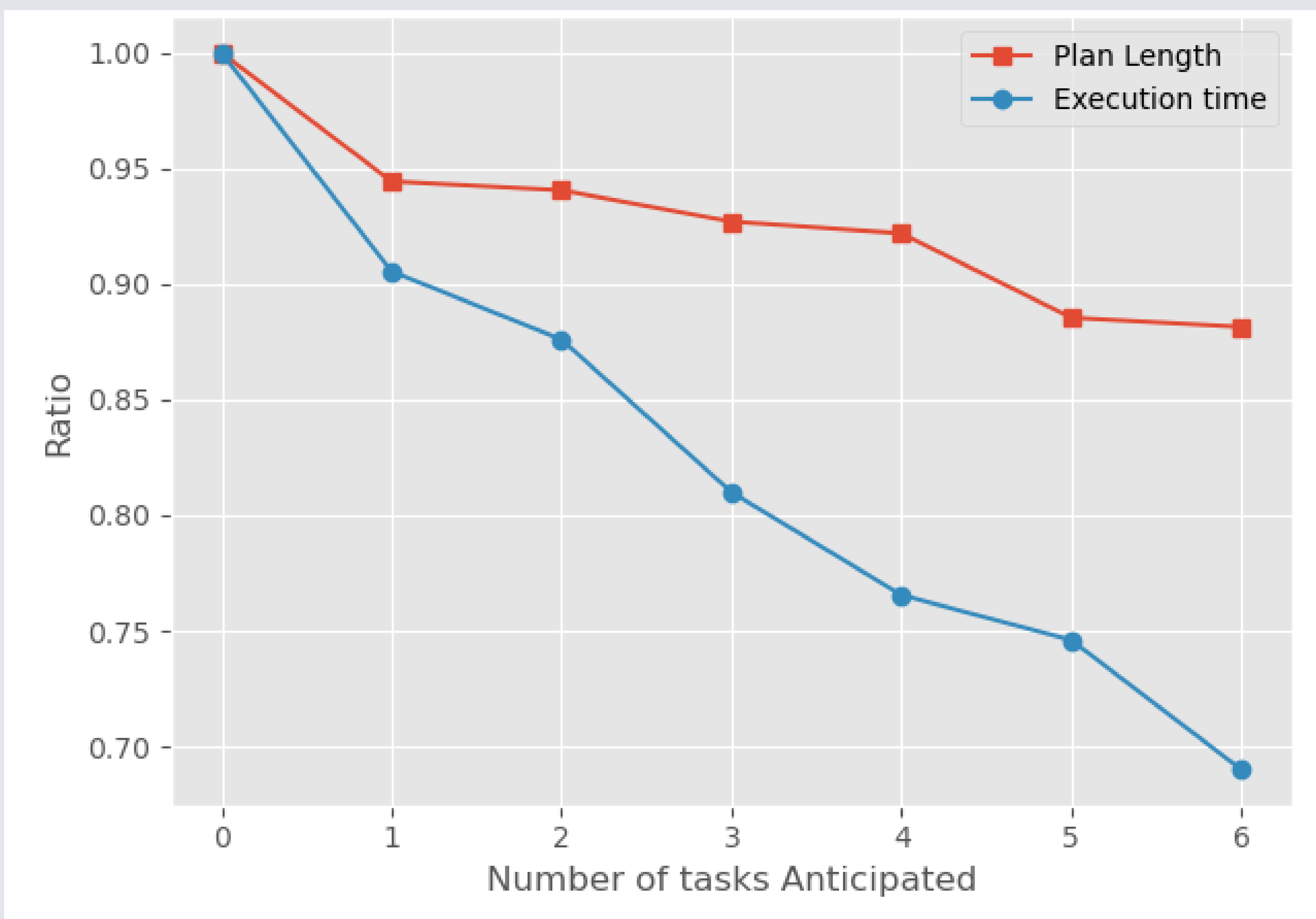
Results

LLMs	Without context			With Context		
	Miss Ratio (Miss.) ⇓	Partial Ordering Count (POC) ⇑	$KRCC$ ⇑	Miss Ratio (Miss.) ⇓	Partial Ordering Count (POC) ⇑	$KRCC$ ⇑
PaLM	0.361	0.974	0.993	0.034	0.994	0.996
GPT-3.5-turbo	0.282	0.676	0.906	0.0698	0.806	0.976
GPT-4	**0.037**	**0.960**	**0.995**	**0.0006**	**1.0**	**1.0**

$KRCC = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}}$

Performance of LLMs for Task Anticipation. Results over 500 experiments with ≈ 20 tasks per experiment.





Mean execution cost ratio and plan length ratio WRT Myopic Agent

Discussion

We describe a framework: **Anticipate&Act**, for task anticipation and action execution by an agent in complex household environments. We use Planning Domain Definition Language (PDDL) as the action language to create a household domain and use the Fast Downward solver to compute plans for any goal state.

We present a 31% reduction in execution time and a 12% reduction in plan length compared to a system that does not anticipate upcoming tasks

Qualitative evaluation in VirtualHome simulation
(Pickup of multiple items for anticipated tasks)

References

1. Dhakal, R., Hossain Talukder, M. R., & Stein, G. J. (2023). Anticipatory Planning: Improving Long-Lived Planning by Estimating Expected Cost of Future Tasks. In 2023 IEEE International Conference on Robotics and Automation (ICRA), 11538–11545.
2. Valmeekam, K., Marquez, M., Olmo, A., Sreedharan, S., & Kambhampati, S. (2023). PlanBench: An Extensible Benchmark for Evaluating Large Language Models on Planning and Reasoning about Change. In Advances in Neural Information Processing Systems (Vol. 36, pp. 38975–38987).
3. McDermott, Drew, et al. "PDDL-the planning domain definition language." (1998).